

# IUPAC DIVISION (VIII) OF CHEMICAL NOMENCLATURE AND STRUCTURE REPRESENTATION

## Report to IUPAC Bureau, August 2004

The following was written before the meetings of the Division Committee and various project groups in Budapest during the period August 29th to September 2nd 2004. Any supplementary information arising from these meetings will be presented verbally or as a supplementary written report in Bled.

### I. Highlights

- I.1 IUPAC-NIST Chemical Identifier.** Imminent release of version 1.0 and good prospects for use by the community (item II.3.1)
- I.2 Organic Preferred Names.** New Blue Book draft completed, including recommendations for selecting Preferred IUPAC Names (PINs), about to be submitted for ICTNS and public review (item II.3.2)
- I.3 Revised Red Book.** Public and ICTNS review completed and final revision in progress (item II.3.4)
- I.4 Graphical representation standards for chemical structures.** A major new project.

### II. Report of activities 2003-2004

**II.1 Context.** The work of the Division of Chemical Nomenclature and Structure Representation is concerned entirely with standards for the transmission of chemical information, and as such addresses the following long-range IUPAC goals:

*b) IUPAC will facilitate the advancement of research in the chemical sciences through the tools that it provides for international standardization and scientific discussion.*

*d) IUPAC will foster communication among individual chemists and scientific organizations, with special emphasis on the needs of chemists in developing countries.*

**II.2 Constitution of the Division.** Division VIII consists of a Division Committee (twelve members from eight nations and nine National Representatives) and an Advisory Subcommittee (39 people from 17 nations). The Joint IUPAC-IUBMB Commission on Biochemical Nomenclature (nine members from six nations) reports to the Division Committee. Division members have access to a Web Discussion Board, on which minutes of meetings and current draft recommendations are posted for comment.

#### **II.3 Current projects**

##### **II.3.1 IUPAC-NIST Chemical Identifier.**

A second test version of the Identifier was distributed in mid-2003; this version extends the range of applicability from discrete covalent organic structures to include organometallic and inorganic structures. A meeting of the project team and representatives of other interested organisations took place at NIST in November 2003, in conjunction with a meeting on CPEP's

XML Data Dictionaries project, to receive feedback and to define the way forward. No significant problems were encountered and a final test version was distributed in July 2004. To recognise the substantial contribution of NIST to the project, the name was changed to the IUPAC-NIST Chemical Identifier (INChI). It is planned to issue version 1.0 of the software in autumn 2004. Consideration will then be given to a further extension of the program to include polymers, and to ways of encouraging use of the new identifier. It has already been incorporated into Chemical Markup Language, and is to be used in the recently initiated PubChem database project of the US National Institutes of Health, and in the Chemical Entities of Biological Interest (ChEBI) database of the European Bioinformatics Institute. The European Patent Office is considering its use as their standard for structure representation. It has been incorporated into the freely available structure-drawing program ChemSketch (from Advanced Chemistry Development).

### ***II.3.2 Organic Preferred IUPAC Names (PINs)***

Work on the new Nomenclature of Organic Chemistry (IUPAC Blue Book), including recommendations for identifying IUPAC-preferred names, is approaching completion. Preparation of a revised draft (more than 1200 pages) following expert review is almost complete, and this version is about to be submitted for ICTNS and public review. Publication is expected in 2005.

### ***II.3.3 Corrections to Revised Nomenclature of Organic Chemistry Section F: Natural Products***

A number of errors and inconsistencies in this document have been pointed out and a substantial Corrections and Modifications document has been published both in print (in PAC) and on the web. The changes have been incorporated into the web version of Revised Section F, with links to the original text.

### ***II.3.4 Revision of “Nomenclature of Inorganic Chemistry”***

Revision of the IUPAC Red Book is approaching completion; the period for ICTNS and public review ended on August 31st 2004, and a final meeting of the project group took place in Budapest in September, to review comments received. Publication in 2005 is expected.

### ***II.3.5 Rotaxanes***

A second draft of recommendations for naming discrete (as opposed to macromolecular) rotaxanes was reviewed by the project team in Budapest on September 1st 2004.

### ***II.3.6 Fullerene nomenclature Part II***

This extension of the published Part I recommendations to larger and more complicated molecules is approaching completion. The period for public and ICTNS review ended on August 31st 2004, and the project team met in Budapest on September 2nd to review comments.

### **II.3.7 Extension of rules for stereodescriptors to include coordination numbers 7-12**

New project now in progress.

### **II.3.8 Nomenclature of cyclic peptides**

This old project of the Joint IUPAC-IUBMB Commission on Biochemical Nomenclature has now been resuscitated. Expert review of draft recommendations was completed at the end of July 2004, and the document will now be submitted to ICTNS and public review.

### **II.3.9 Graphical representation standards for chemical structures**

This is a new project arising out of a scoping exercise carried out in 2002-2003. Various draft recommendations are under consideration (in particular proposals for representation of stereochemistry), and the project team met for the first time in Budapest on September 1st and 2nd.

#### ***Macromolecular Nomenclature Projects***

The following five projects are managed in full cooperation with Division IV's Subcommittee on Macromolecular Terminology:

### **II.3.10 Structure-based nomenclature for cyclic macromolecules**

This is expected to be ready for review in October 2004.

### **II.3.11 Nomenclature for macromolecular rotaxanes**

This has been put on hold pending completion of the rotaxanes project (see item II.3.5)

### **II.3.12 Nomenclature of dendrimers and hyperbranched oligomers and polymers**

This is expected to be ready for review in October 2004.

### **II.3.13 Process-based nomenclature for chemically modified polymers**

This project has been redefined as Source-based Nomenclature for Modified Polymer Molecules and a proposal for project extension will be submitted.

### **II.3.14 Source-based nomenclature of single-strand organic polymers**

This is a new project; it is now proposed to change the title to the more explicit Source-based Nomenclature of Organic Homopolymers and Copolymers.

## **II.4 *Joint IUPAC-IUBMB Commission on Biochemical Nomenclature (JCBN)***

Joint activities of the two Unions are channelled through this Joint Commission, which has been reconstituted in the light of the changes in arrangements for managing IUPAC's work. The main activities of the Commission are

### **II.4.1 Maintenance and updating of the Enzyme List**

This is a very substantial and continuous operation drawing on advice from IUPAC participants on chemical names for substrates, reagents and products.

#### **II.4.2 Maintenance and development of specialised naming systems for natural products**

Classes of natural product of interest to biochemists, especially steroids, amino acids and peptides, carbohydrates, lipids and nucleic acids require local specialised systems for naming. A project to review and update carbohydrate nomenclature is being assembled.

#### **II.4.3 Advice for biochemists on names for specific compounds of biochemical importance**

The development of a synonyms database for compounds in common biochemical usage is under consideration, building on work in progress at the European Bioinformatics Institute and at the University of Missouri.

### **II.5 Scoping Exercises**

#### **II.5.1 Preferred IUPAC names for inorganic compounds**

A proto-project team met on September 29th in Budapest to consider approaches to project development in this area.

### **III. Possible future activities**

At the Division Committee meeting on August 30th and 31st in Budapest, the following were reviewed as desirable areas for future activity:

- Nomenclature of metallacycles
- Preferred structure-based names for macromolecules (consequences of Blue Book revision)
- Stereochemistry (interdisciplinary book)
- Inorganic nanomaterials (arising from discussion with the US Centre for Biological and Environmental Nanotechnology)
- Boron nomenclature (revision and extension)
- Lipid nomenclature (revision)

*Alan McNaught*  
August 2004